

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

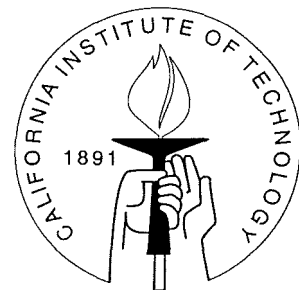
PASADENA, CALIFORNIA 91125

THE EFFECTS OF PAYOFF MAGNITUDE AND HETEROGENEITY ON
BEHAVIOR IN 2×2 GAMES WITH UNIQUE MIXED STRATEGY
EQUILIBRIA

Richard D. McKelvey

Thomas R. Palfrey

Roberto A. Weber



SOCIAL SCIENCE WORKING PAPER 991

February 1999

The Effects of Payoff Magnitude and Heterogeneity on Behavior in 2×2 Games with Unique Mixed Strategy Equilibria

Richard D. McKelvey

Thomas R. Palfrey

Roberto A. Weber

Abstract

The Logit version of Quantal Response Equilibrium predicts that equilibrium behavior in games will vary systematically with payoff magnitudes, if all other factors are held constant (including the Nash equilibria of the game). We explore this in the context of a set of asymmetric 2×2 games with unique totally mixed strategy equilibria. The data provide little support for the payoff magnitude predictions of the Logit Equilibrium model. We extend the theoretical Quantal Response Equilibrium model to allow for heterogeneity, and find that the data fit the heterogeneous version of the theory significantly better.

The Effects of Payoff Magnitude and Heterogeneity on Behavior in 2×2 Games with Unique Mixed Strategy Equilibria

Richard D. McKelvey

Thomas R. Palfrey

Roberto A. Weber

1 Introduction

Most research in experimental economics relies implicitly on an assumption that marginal payoff incentives in the experimental environment are sufficiently great to successfully induce preferences and thereby provide a good approximation of the theoretical economic environment being studied. On the other hand, it has been argued that there are other behavioral components that may contaminate experiments where incentives are not high enough¹. This contamination, it is argued, generally takes either the form of “random noise” in the data or the form of systematic departures due to other competing and uncontrolled incentives such as social utility. Until recently, there has been no satisfactory predictive model to apply to this issue of payoff magnitude. There has been some success accounting for deviations from standard theory by appealing to random noise and social utility, and these accounts are often qualitatively consistent with payoff magnitude explanations.² There have also been experimental studies of payoff magnitude effects in both economics and psychology³, but very few game theory experiments have been designed and run to test and calibrate specific quantitative models of payoff magnitude effects.

Quantal Response Equilibrium (QRE) [McKelvey and Palfrey (1995)] provides a model that makes specific predictions about how equilibrium behavior in games should change in response to payoff magnitudes. According to this theory, there are two effects of increased payoff magnitude in games. The first effect is direct, and is present in many models of payoff magnitude borrowed from the stochastic choice literature: increasing the magnitude of incentives will reduce decision errors by subjects. They will make optimal choices with greater frequency. The second effect is an indirect general equilibrium effect. In strategic situations the decision errors of one subject will affect the payoffs of the other subject. This in turn will affect the choice frequencies of that subject which

¹See for example Harrison (1989).

²See, for example, McKelvey and Palfrey (1992), Palfrey and Prisbrey (1997), and Cooper, DeJong, Forsythe, and Ross (1996).

³See Grether and Plott (1979) and Siegel (1961).

then feed back on the original subject. An equilibrium with these decision errors is a quantal response equilibrium. Thus changes in payoff magnitudes can change the quantal response equilibrium of the game, even if there is no change in the standard (Nash) equilibrium of the game.

This paper presents the results from an experiment, using a collection of very simple 2×2 games of complete information, similar to those originally studied by Ochs (1995). The games we concentrate on in our study are interesting in that they all possess the same unique mixed strategy Nash equilibrium for either one or both players. This is due to the fact that they only differ in payoff magnitude, a difference to which the Nash equilibrium concept is insensitive. On the other hand, the Logit version of QRE makes different predictions for the different games. Maximum likelihood estimates are used to test between these equilibrium notions, random play, and a model which looks at a hybrid between Nash equilibrium and random play, which we call Noisy Nash Equilibrium (NNE). We find problems with all of these models, but find that QRE fits the data best.

We conjecture that one of the problems in fitting the data is that these models assume that all agents have identical error rates, or rationality parameters. In nearly all attempts to carefully measure subject heterogeneity in economics and game theory experiments, significant and potentially important sources of heterogeneity have been found⁴. In the last section, we develop a heterogeneous extension of the Logit QRE model. We find that this heterogeneous subject model generates substantially different predictions and produces a significant improvement in fit.

2 Payoff Magnitude Effects and Quantal Response Equilibria

The Quantal Response model incorporates error into the best response of players in a game. Thus, the perfectly rational model of choice usually assumed to govern players' actions is replaced by a probabilistic one where better responses are more likely to be played, but no action is played with certainty.

The logistic specification of the QRE, which we use here, measures error in terms of a precision parameter, λ , which is inversely related to the variance of the error.⁵ If the expected utility to strategy j for agent i is denoted by u_{ij} then the probability that i will

⁴See, for example, Cox, Smith, and Walker (1983), El-Gamal and Grether (1995), Palfrey and Prisbrey (1997), and Stahl and Wilson (1995). While these papers investigate different possible sources of heterogeneity than the one we look at here, in all cases the heterogeneous models significantly outperform the homogeneous models.

⁵For a formal and more detailed explanation of QRE, see McKelvey and Palfrey (1995, 1996, 1998). Those papers also provide discussion about the conceptual interpretation of λ .

use strategy j is given by the logit formula:

$$p_{ij} = \frac{e^{\lambda u_{ij}}}{\sum_k e^{\lambda u_{ik}}},$$

where k runs over the available strategies for agent i . If $\lambda = 0$, then players are acting entirely in error (i.e. completely unresponsive to payoffs), while for $\lambda = \infty$, players' actions are equivalent to perfect expected utility maximizing behavior.

	Game A		Game B		Game C		Game D	
	L	R	L	R	L	R	L	R
U	9,0	0,1	9,0	0,4	36,0	0,4	4,0	0,1
D	0,1	1,0	0,4	1,0	0,4	4,0	0,1	1,0

Table 1. Payoff Tables for Games A-D

Table 1 presents the games studied in this paper. Game A is the same as a payoff matrix recently studied by Ochs (1995). Games B and C are variations of Game A which only involve payoff magnitude changes. In game B, the column player's payoffs of Game A are multiplied by a factor of four. In Game C, both the row and column player's payoffs of Game A are multiplied by a factor of 4. Game D was included primarily as a replication of Ochs (1995) study.

All four games have unique mixed strategy Nash equilibria. Since Nash equilibria are insensitive to positive affine transformations in the payoffs, they are identical for Games A, B, and C. Letting p denote the probability the row player chooses U and q denote the probability the column player chooses L , these equilibrium probabilities are $p^* = 0.5$ and $q^* = 0.1$. For Game D, the Nash equilibrium is at $p^* = 0.5$ and $q^* = 0.2$. Thus the predictions of Nash equilibrium are easily summarized⁶: (i) in all games, the row players' choice behavior should be completely random; (ii) in Games A, B, and C, the column players should choose L 10% of the time; (iii) In Game D, the column player should choose L 20% of the time.

In contrast with Nash equilibrium, QRE is sensitive to differences in payoff magnitude and makes different predictions for the games. Figure 1 represents the QRE as a function of λ for Games A, B, and C, while Figure 2 shows the QRE for Games A and D. Here, $p_X = p_X(\lambda)$ and $q_X = q_X(\lambda)$ for $X \in \{A, B, C, D\}$, denote the probability of choosing U and L for any given value of λ . In each case, the leftmost points in the graph correspond to the randomness predicted by low values of λ , while the probabilities corresponding

⁶This actually glosses over a subtle issue of multiple equilibria that arises due to the experimental protocol. Consider an experiment with 16 subjects (8 row and 8 column), where subjects are randomly repaired with a new opponent every time they replay the game. This creates many new equilibria. For example, having 4 of the row players always play U and the other 4 row players always play D is not inconsistent with equilibrium, since, from the column players' point of view, this is indistinguishable from all eight of the players independently randomizing 50/50 between U and D . There are a continuum of multiple equilibria of this sort. However, aggregate choice frequencies are the same in all such equilibria.

to high values of λ are, in the limit, equivalent to the Nash equilibrium prediction. The pattern of convergence is very interesting in these games. Specifically, the column player's QRE strategy always begins at $(.5, .5)$ and converges monotonically to the Nash equilibrium. But the row players' QRE strategy actually begins at the Nash equilibrium (since it is $(.5, .5)$), then progressively overplays (relative to Nash equilibrium) the U strategy until eventually converging back to $(.5, .5)$.

The two features that vary across the QRE graphs of the three games are the speed with which convergence takes place and how much the row player overplays U . In Figure 1, it can be seen that convergence of the QRE to the Nash equilibrium occurs most rapidly in Game C, which has the highest payoffs, second most rapidly in Game B, and slowest in Game A. Furthermore, the QRE predicts that, for intermediate values of λ , the row player will be overplay strategy U more severely in Games A and C relative to the Nash prediction than in Game B.

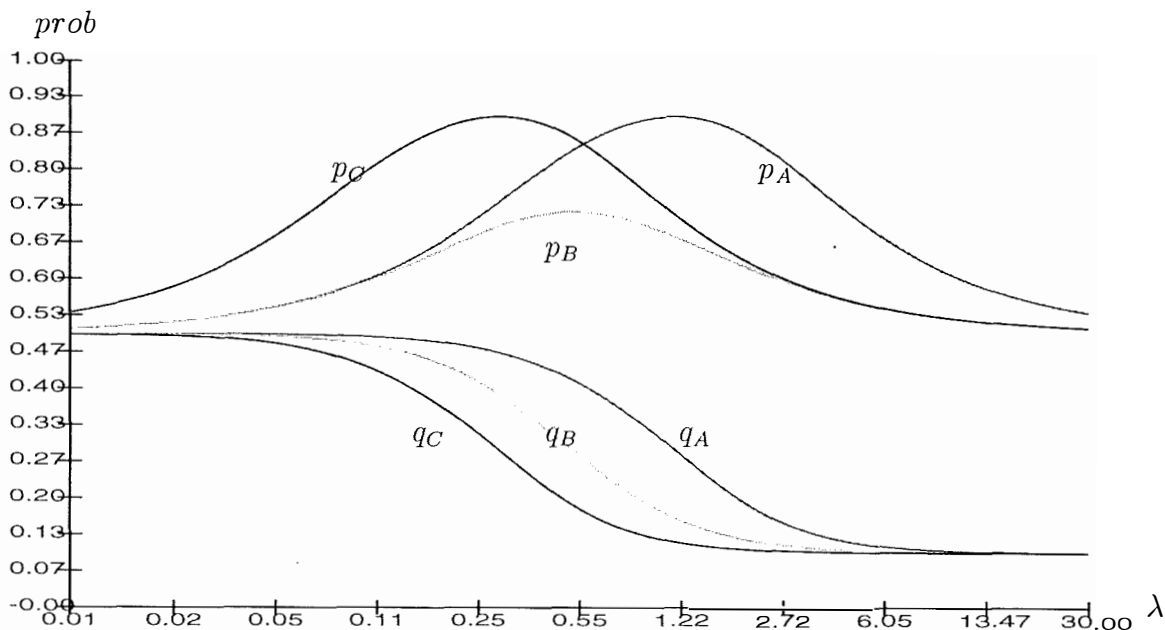


Figure 1: QRE correspondence for Games A, B, and C

In our most basic model of homogeneous subjects, we assume that λ is exogenously determined, and is identical for all subjects, and identical across games. This yields some specific predictions about differences in behavior in these games: For the first three games, represented in Figure 1, one can readily see that for any value of λ , this implies for the row player, $.5 = p^* < p_B < p_A$, and $.5 = p^* < p_B < p_C$, and for the column player $.5 > q_A > q_B > q_C > q^* = .1$. Similarly, for the second subset of games (Game A and Game D) represented in Figure 2, we predict for the row player, $.5 = p^* < p_D < p_A$, and for the column player we expect $.5 > q_D > q_A > q^* = .1$. The experiments of this paper

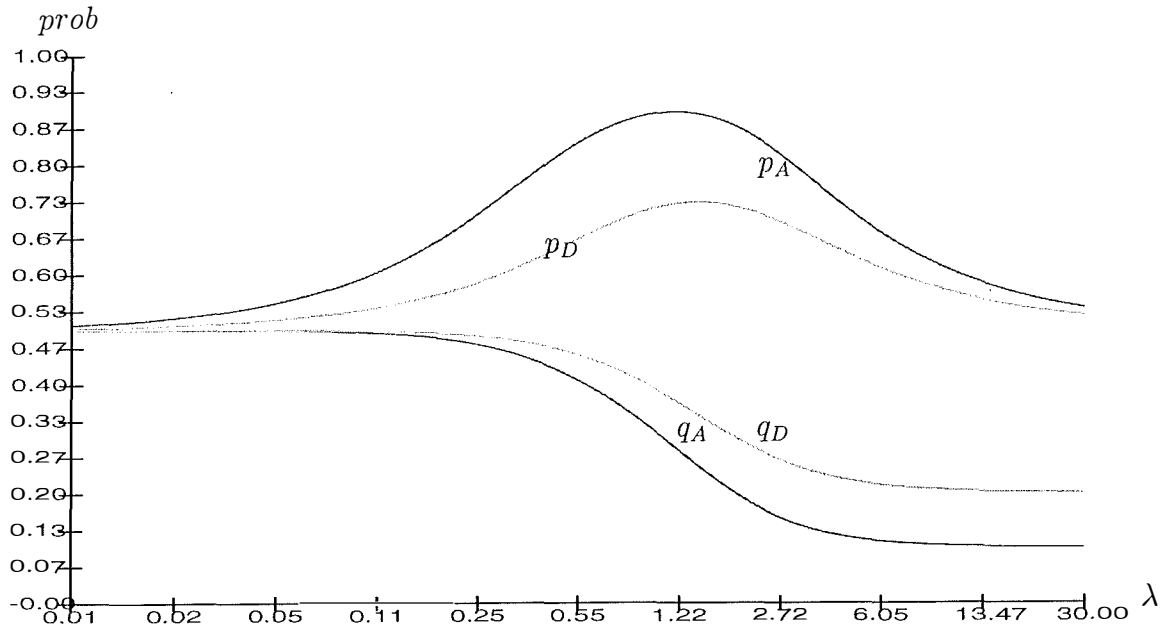


Figure 2: QRE correspondence for games A and D

were designed to test these predictions.

In order to provide an alternative model of behavior we compare the fit of QRE to that of a simple one parameter model, which adds error to the Nash model in a different way, similar to the approach of Smith and Walker (1993). Like the QRE, this model also spans a range of behavior for which the Nash equilibrium and the random model are the extremes. But instead of embodying an equilibrium restriction like QRE, this model simply assumes that players play Nash with probability γ and play randomly with probability $(1 - \gamma)$. Therefore, the range of γ is the interval $[0, 1]$ and, for a given value of γ , the equilibrium correspondence is: $\hat{p} = (1 - \gamma).5 + (\gamma)p^*$ and $\hat{q} = (1 - \gamma).5 + (\gamma)q^*$. We call this the *Noisy Nash Model* (NNM). It is similar to the QRE model, in the sense that when $\gamma = 0$ the model predicts random play, and for larger values of γ , the equilibrium approximates the Nash prediction. However, it differs from the QRE in that it is not a full equilibrium model. In particular, while subjects choose actions with some error (which is equal across actions and is a function of γ) they do not take into account the fact that other players are also choosing with error. This implies that NNM generally makes different predictions from QRE. For example, NNM predictions do not vary across Games A, B, and C, which all have the same unique Nash equilibrium. Also, for all games NNM predicts that the row player always chooses up with probability .5 independent of γ . That is, $\hat{p}(\gamma) = .5$ for all $\gamma \in [0, 1]$.

3 Experimental Design

We conducted a total of eight experimental sessions on the above games, using as subjects undergraduate and graduate students at the California Institute of Technology with little or no formal training in game theory. Each session used twelve subjects, who participated in two of the above games. Each game in a session was played 50 times by each subject who was anonymously matched with a subject of the opposite type in each period, using a random matching procedure. No two subjects were paired together twice in a row, and subjects were fully informed of the matching procedures. The experiments were conducted through computers by which subjects were presented with the complete payoff matrix and made their choices by clicking on the appropriate row or column with a mouse. All subjects could observe, on their computer screen, the full history of actions and payoffs for pairings in which they had participated.

Of the above four games, A, B, and C were played twice with each of the other two, once before and once after the other game. Table 2 summarizes the treatments in each session. Note that Game D was only played with Game A. For every pair of games that was run, we conducted two sessions in order to control for sequencing effects. For example, session 1 is an *AB* session while session 2 is a *BA* session.

Session	First Game	Second Game
1	A	B
2	B	A
3	B	C
4	C	B
5	A	C
6	C	A
7	A	D
8	D	A

Table 2. First and Second Treatments by Experimental Session

Before the experiment, each of the 12 subjects was randomly assigned to either the row or column position, and these roles did not change for the entire experiment. Participants were informed that each unit of payoff represented \$0.10, and that they would be paid this amount, in cash, at the end of the experiment. Prior to the experiment, subjects were read instructions⁷ and guided through four instructional periods, where all subjects observed each outcome cell once.

⁷The instructions are in the appendix.

At the end of the experiment, subjects were privately paid their earnings, in addition to a \$5.00 participation bonus. Earnings varied substantially by treatment and player type. The average earnings for column players were \$2.90 for Game A, \$11.18 for Game B, \$11.72 for Game C, and \$2.44 for Game D. For the row players, the average earnings in each game were \$8.17, \$8.07, \$31.22, and \$5.48, respectively, for Games A, B, C, and D. The substantially lower payoffs for Games A and D, which were played together only in Sessions 7 and 8, led us to omit the announcement of the amount of the participation bonus until after the experiment in these sessions, at which time the bonus was announced to be \$10.00 instead of \$5.00.

4 Results

Table 3 provides a summary of the results for each session, as well as aggregate results for each of the games. At this very high level of aggregation, the effects of varying payoff magnitude do not appear to be particularly strong and therefore provide at most weak support of the hypotheses. Increasing the column players' payoffs from Game A to Game B has little effect on the behavior of those players, but it appears to make the row players somewhat less likely to play action U . The latter is predicted by the payoff magnitude hypotheses. Increasing the row players payoffs in Game C appears to have the effect of decreasing the frequency of the action U and decreasing that of action L . The latter effect is consistent with the hypothesis, but the former one is not. The effects of altering payoff magnitude vary within the individual sessions. Finally, notice that for none of the games is it the case that the observed behavior corresponds to that predicted by the Nash equilibrium. In all games, U and L are overplayed relative to the Nash equilibrium. This replicates the findings in Ochs (1995) and is consistent with the predictions of QRE.⁸

⁸An alternative hypothesis for this observation is that (U, L) is some kind of focal point. While our design is not equipped to reject this hypothesis, it seems unlikely to be an driving force in the data, given the absence of a multiple equilibrium problem.

Session	Game	U	L	n
1	A	0.680	0.317	300
	B	0.573	0.197	300
2	B	0.647	0.237	300
	A	0.623	0.243	300
3	B	0.607	0.363	300
	C	0.570	0.393	300
4	C	0.623	0.163	300
	B	0.693	0.180	300
5	A	0.623	0.187	300
	C	0.590	0.197	300
6	C	0.593	0.273	300
	A	0.607	0.223	300
7	A	0.640	0.230	300
	D	0.457	0.313	300
8	D	0.643	0.343	300
	A	0.683	0.243	300
Aggregate	A	0.643	0.241	1800
	B	0.630	0.244	1200
	C	0.594	0.257	1200
	D	0.550	0.328	600

Table 3. Experimental Results by Session and Game

We use a standard χ_2 test to test the null hypothesis that the observed actions are independent of the game being played. The observation that the row players select D more frequently in Game C is statistically significant at the 0.01 level in a comparison to Game A and at the 0.10 level in a comparison to Game B. Furthermore, the differences between Games A and D in the observed frequency choices for both players is significant at the 0.001 level, which is not surprising for the column player, since the Nash predicted frequency of choice L is different between the two games.⁹ This is not true, however, for the row player's actions, which Nash equilibrium predicts to be the same across the games.

Table 4 presents the aggregate results for each game, as well as the estimates of λ . In the tables, p and q represent the observed frequencies of the row player choosing U and of the column player choosing L , respectively. Maximum likelihood estimation was used to obtain an estimate of λ for these frequencies. The values λ_{LO} and λ_{HI} provide a 95% confidence interval for this estimate and \hat{p} and \hat{q} represent the corresponding QRE probabilities. QRE, NASH, and RAND give the log-likelihoods of the maximum likelihood

⁹Our reported significance levels assume that observations are i. i. d. Since, as we demonstrate later, there are important sources of dependence in the data due to heterogeneity and since there appear to be important cohort effects, these levels overstate the statistical significance of these findings.

estimator of λ , the Nash equilibrium solution for that game ($\lambda = \infty$), and of random play ($\lambda = 0$), respectively.

For all four games, a likelihood ratio test rejects, at the 0.001 level, the hypothesis that the observed data is consistent with random play (where $\lambda = 0$ and all strategies are played with equal probabilities). Similarly, for all four games, the hypothesis that the experimental results are consistent with Nash equilibrium play can be rejected at the 0.001 level.¹⁰

Game	\hat{p}	p	\hat{q}	q	λ	λ_{LO}	λ_{HI}	QRE	NASH	RAND
A	0.690	0.643	0.115	0.241	5.38	4.73	6.25	-2286.1	-2388.7	-2495.3
B	0.711	0.630	0.220	0.244	0.75	0.64	0.89	-1478.0	-1602.0	-1663.5
C	0.635	0.594	0.107	0.257	1.97	1.58	2.63	-1603.8	-1634.9	-1663.5
D	0.590	0.550	0.210	0.328	7.33	4.41	18.13	-817.3	-822.9	-831.8

Table 4. Summary of Results and Estimates for all Games

In all four games, subjects in the row position overplayed strategy U relative to the Nash equilibrium prediction, particularly in later periods¹¹. At the same time, column players overplayed strategy L . Furthermore, in Games A and C the QRE predicts that players will over play U more than in the other two games, for intermediate values of λ , and we find that this is true for game A, although not to the extent that the model predicts.

Figures 3 through 6 again provide the QRE as a function of λ for each of the games, aggregated over all data of that experiment. In addition, the experimental frequencies for each subset of 10 periods are plotted according to their estimated¹² value of λ . Periods 1-10, 11-20, 21-30, 31-40, and 41-50 are represented by 1, 2, 3, 4, and, 5, respectively, in the figure. The aggregated data for that game, or periods 1-50, are indicated by the number 6. We find no systematic trends in the estimated values of λ over time.¹³

¹⁰If λ represents an exogenously determined measure of precision (or inverse error) in the population, then it follows that the estimated values of this parameter should not differ between games. We tested this assumption by estimating a model in which the parameter is constrained to have the same value across games ($\lambda_A = \lambda_B = \lambda_C = \lambda_D$). We reject the constrained model in favor of the unconstrained model.

¹¹In fact, given the actual column strategy frequencies, U gives the row player a considerably higher expected payoff than L .

¹²Each 10-period block was estimated separately.

¹³In order to measure whether the parameter λ captures any learning that may take place across similar games, we recomputed the parameter estimates separately for each game when it was the first game played in a session and when it was the second (see Table 2). We found that for games A, B, and D, the estimated parameter values were higher when the game was played second than when it was played first. This difference was significant for games B and D. For game C, however, the difference was significant and in the opposite direction. This analysis of the data, therefore, provides weak evidence at best that values of λ are increasing with experience.

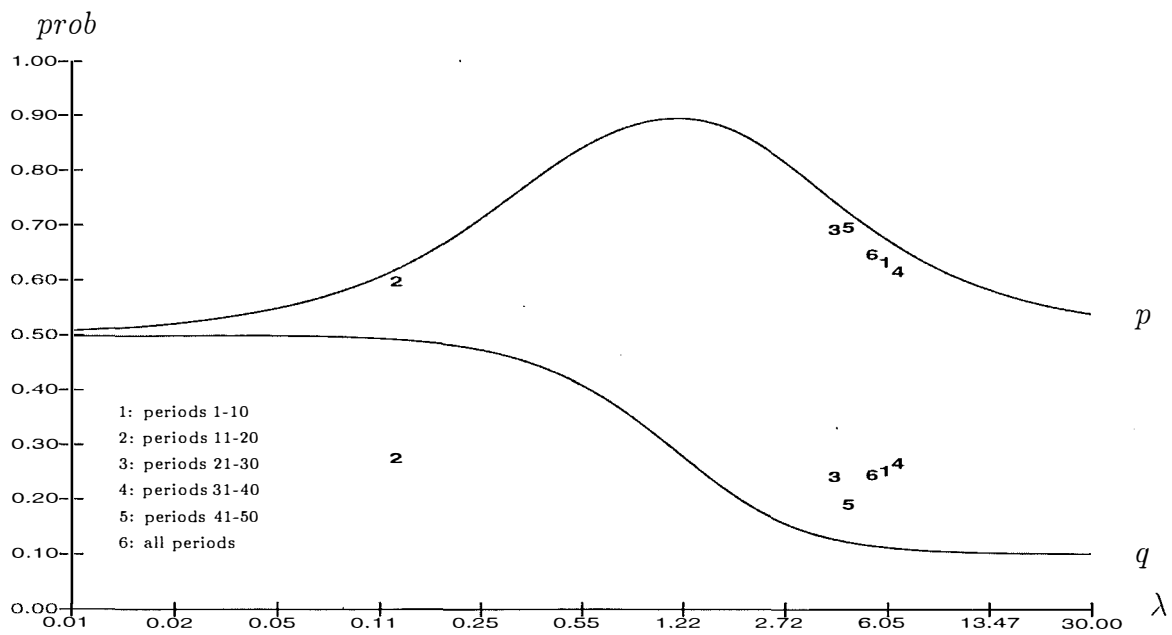


Figure 3: Data for Game A

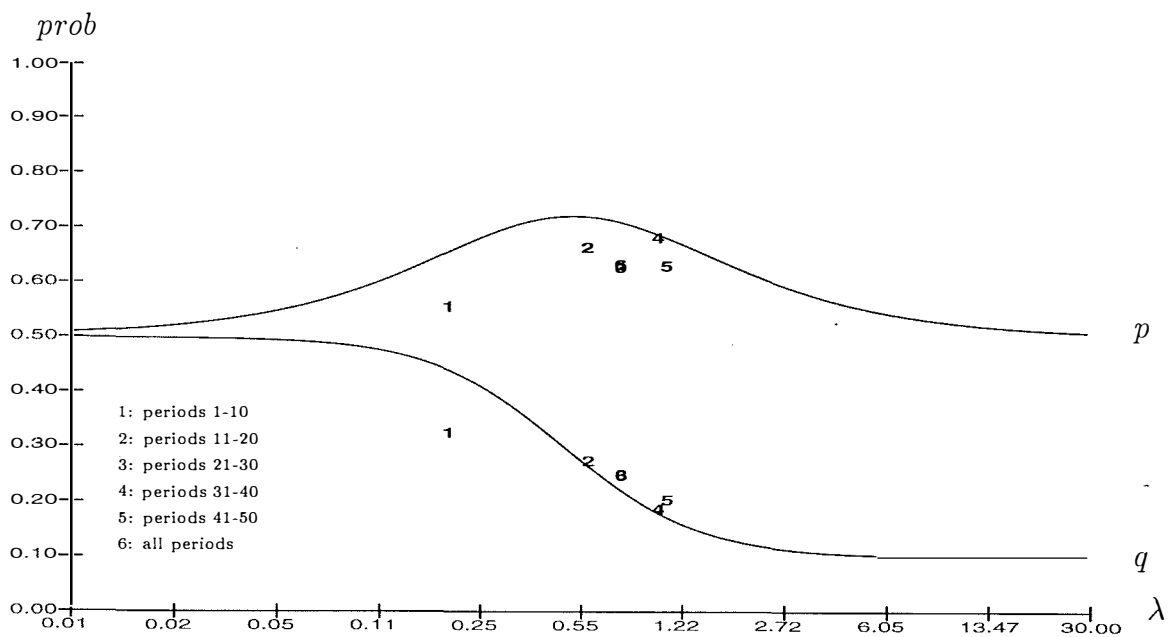


Figure 4: Data for Game B

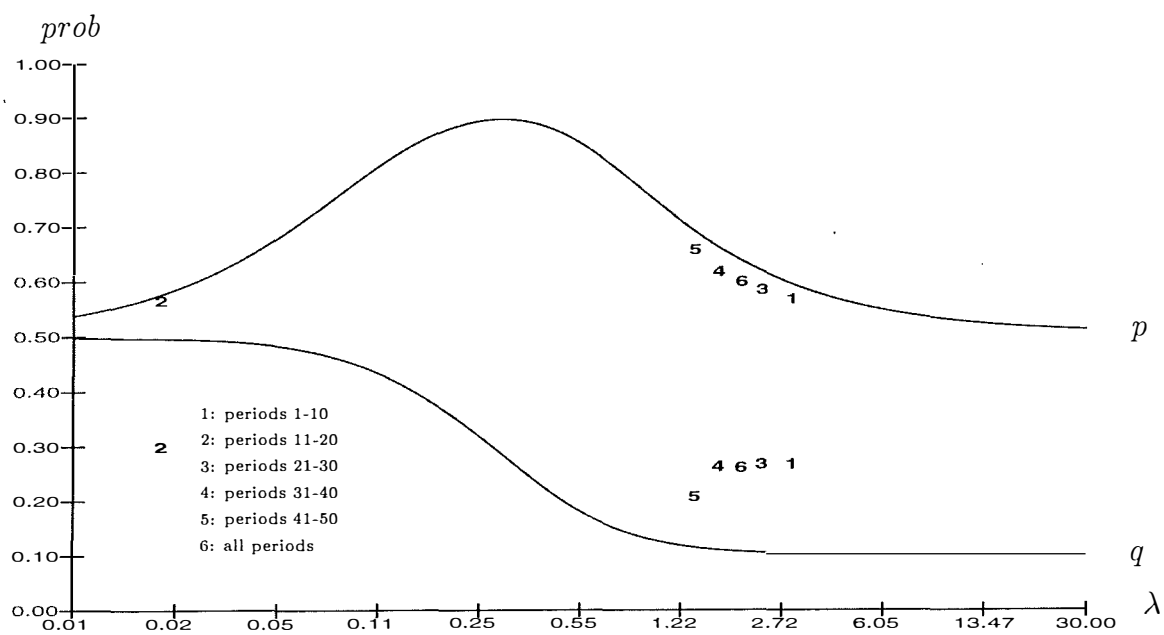


Figure 5: Data for Game C

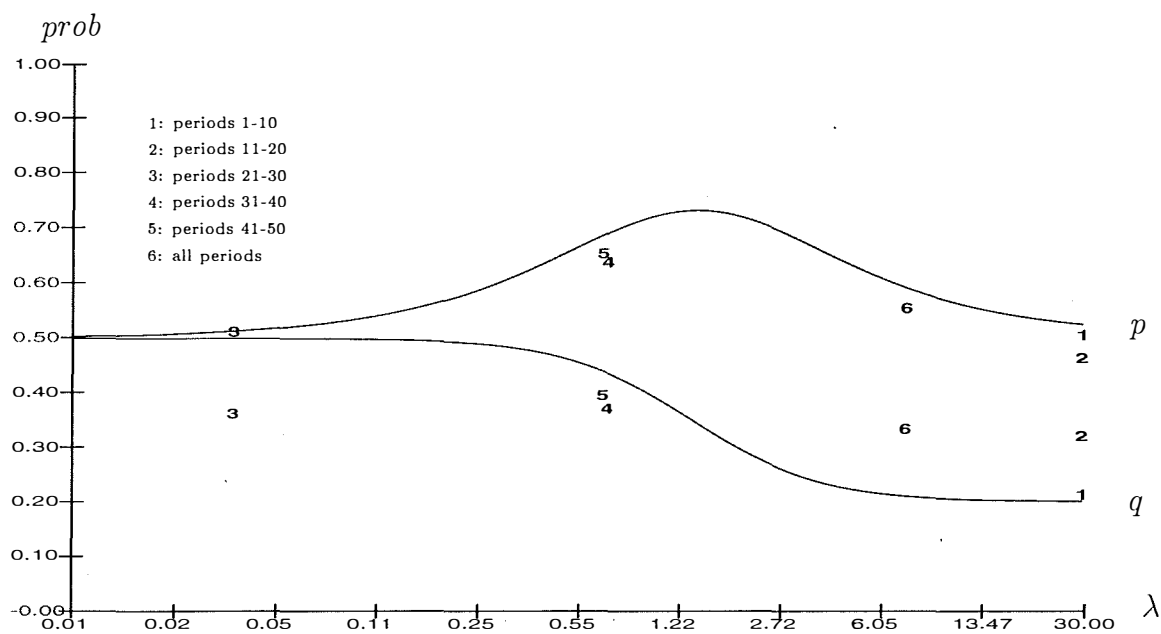


Figure 6: Data for Game D

In these figures, it can again be seen that subjects regularly overplayed U and L relative to the Nash equilibrium prediction but that these observed results fit the QRE model only slightly better. It is interesting to note that the estimates of λ do not register at intermediate values of λ , especially in games A and C, where the QRE predicts high probabilities of U for the row player.

While the QRE clearly outperforms both the Nash model and the random model, one could argue that these alternatives do not provide a very strong test of the theory, since they are extreme points of the QRE correspondence. Therefore, we also test the Noisy Nash Model (NNM) discussed above.

Table 5 presents the aggregate results for each game of the estimation of the NNM. For all games, likelihood ratio tests reject both the Nash prediction and random play in favor of NNM, at all reasonable significance levels. In addition, for Games A, C, and D, NNM provides a better fit for the data than does the QRE. For Game B, the opposite is true, with QRE providing a better fit for the Game B data than does NNM.

Game	\hat{p}	p	\hat{q}	q	γ	γ_{LO}	γ_{HI}	NNM	NASH	RAND
Game A	0.500	0.643	0.240	0.241	0.65	0.60	0.69	-2240.8	-2388.7	-2495.3
Game B	0.500	0.630	0.244	0.244	0.64	0.58	0.69	-1498.8	-1602.0	-1663.5
Game C	0.500	0.594	0.256	0.257	0.61	0.55	0.66	-1515.2	-1634.9	-1663.5
Game D	0.500	0.550	0.329	0.328	0.57	0.45	0.69	-795.7	-822.9	-831.8

Table 5. Estimates for NNM for all Four Games

Table 6 gives a summary of the various predictions from the QRE theory (additional predictions follow from transitivity.) The table indicates which predictions are in the correct direction, as well as a z -value for each, which would be distributed as unit normal under the assumption of independence¹⁴. Predictions based on payoff magnitude effects are marked (P). Asterisks next to the corresponding z values indicate which are significant at the .05 level.

For nine of the twelve predictions, the direction is correct, and all but one of those would be significant at the .05 level under the assumption of independence. None of the four predicted payoff magnitude effects are significant at the .05 level. All other predicted effects are significant and have the correct sign.

It is evident from the above analysis that the hypothesized payoff magnitude effects are not borne out in our data. On theoretical grounds, recall that these predictions were only valid under the maintained hypothesis that all players have the same λ . Such heterogeneity would nullify these predictions. Therefore, in the next section we generalize

¹⁴The assumption of independence is not a good one here since the observations consist of multiple observations from a single individual, which would tend to be correlated. Correlation would cause the z -value to have a larger variance than one. We report the z values nevertheless because they are still useful in judging the relative orders of magnitude of significance between different rows of the table

the QRE model to allow for heterogeneity, to see if this can account for the apparent absence of payoff magnitude effects.

QRE Prediction	Actual	Direction correct?	z- value
$p^* < p_B$	$.500 < .630$	Y	9.001*
$p_B < p_A$ (P)	$.630 < .643$	Y	.726
$p_B < p_C$ (P)	$.630 > .594$	N	-1.810
$.5 > q_A$	$.500 > .241$	Y	21.977*
$q_A > q_B$ (P)	$.241 < .244$	N	-.188
$q_B > q_C$ (P)	$.244 < .257$	N	-.735
$q_C > q^*$	$.257 > .100$	Y	18.129*
$p^* < p_D$	$.500 < .550$	Y	2.449*
$p_D < p_A$	$.550 < .643$	Y	4.064*
$.5 > q_D$	$.500 > .328$	Y	8.426*
$q_D > q_A$	$.328 > .241$	Y	4.193*
$q_A > q^*$	$.241 > .100$	Y	19.940*

Table 6. Summary of QRE predictions versus actual behavior under homogeneity

5 Heterogeneity

Examining the individual data, it is apparent that there is some degree of heterogeneity within the subject pool. One way to see this is illustrated in Figure 7. In each row of this figure, there are two graphs. In the left (right) graph the solid line represents the cumulative distribution of the actual frequency of choice of U (L) by the row (column) players. Each observation is an individual row (column) player's frequency of choice of U (L) over the 50 periods of the experiment. The dotted line represents the cumulative distribution of observed histories by the column (row) players of the choice frequencies of their opponents. Here, each observation represents the history of choices by that subject's opponents over the 50 periods of the experiment¹⁵. Since the subjects are randomly matched with a new subject each period, the dotted line represents the sampling distribution that should be observed in the row player frequencies of choice if all row (column) players are acting the same.

Figure 7 shows that in every game, there is more variance in the actual frequencies (solid line) than in the observed, or expected frequencies (dotted line). Thus, the

¹⁵Since subjects are randomly rematched at the beginning of each period, the actions of the "opponent" are really the 50 choices that the players observed. These choices were not made by one player, but rather by the six players of the other type, in the order determined by the matching procedure.

within subject variance in behavior is less than the total variance, indicating that there is between subject variance in behavior. This means that there is heterogeneity in the behavior of the individual subjects. If there were no heterogeneity, then there should be the same amount of variance across histories as there is across actions of the players. An F test for the differences of the variances of the actual versus observed actions shows that the difference is significant at the .01 level for all eight cases¹⁶.

The fact that there is larger variance of actions contingent on opponent's choice than there is on the opponent's choice itself is a strong indicator of heterogeneity. This heterogeneity is also evident when the individual frequencies of responses are compared directly. Therefore, it seems that a model which incorporates this variability of actions into players' choices in the games may provide a better description of actual behavior. Moreover, if heterogeneity is present, the models in the previous section are misspecified and hence need to be corrected by estimating a model which explicitly incorporates heterogeneity.

The exact nature of the misspecification problem can also be seen from inspection of the QRE correspondences and from the figures that compare our fitted estimates to these correspondences (Figures 3-6). Especially in games A and C, the frequency of U choices by row players is never as high as what would be predicted by the QRE given the actions of the column players. Furthermore, at the estimated value of λ , the frequency of L by the column players is too high. However, notice that the equilibrium graph of the row frequencies is close to .50 for both low and high values of λ . Thus, if there is some variance of λ , then it is possible that we could recover a lower *average* estimate of λ (close to the peak of the "UP" curve of the QRE correspondence) and simultaneously overcome both the problem of overestimation of U and the problem of underestimation of L .

Figures 8-9 illustrate this point by displaying the QRE for a fixed variance in the distribution of λ across individuals as a function of the mean. These figures correspond to Figures 1 and 2, when heterogeneity is introduced into the model. Note that the heterogeneity dampens the variation in the QRE, as expected. More importantly, also observe that the qualitative predictions about the ordering of the row and column choice frequencies across games are no longer valid. In fact, there is very little separation of predictions between the various games, which may well explain why we found little support for these qualitative predictions in the previous section.

With the above goal in mind, we develop and test the following parametric model of heterogeneity. In the framework of the QRE model, we allow each player, i , to behave according to some a specific value of λ_i , and i believes that other subjects are the same. These values, however, are allowed to vary between subjects according to a distribution. Hence, the previous QRE estimations are a restricted case of this model where the re-

¹⁶The F values are, for game A ($n = 36$): (row) 5.323, (col) 9.340, for game B ($n = 24$): (row) 11.535, (col) 4.623 for game C ($n = 24$): (row) 15.876, (col) 4.083 and for game D ($n = 12$): (row) 4.567, (col) 21.194

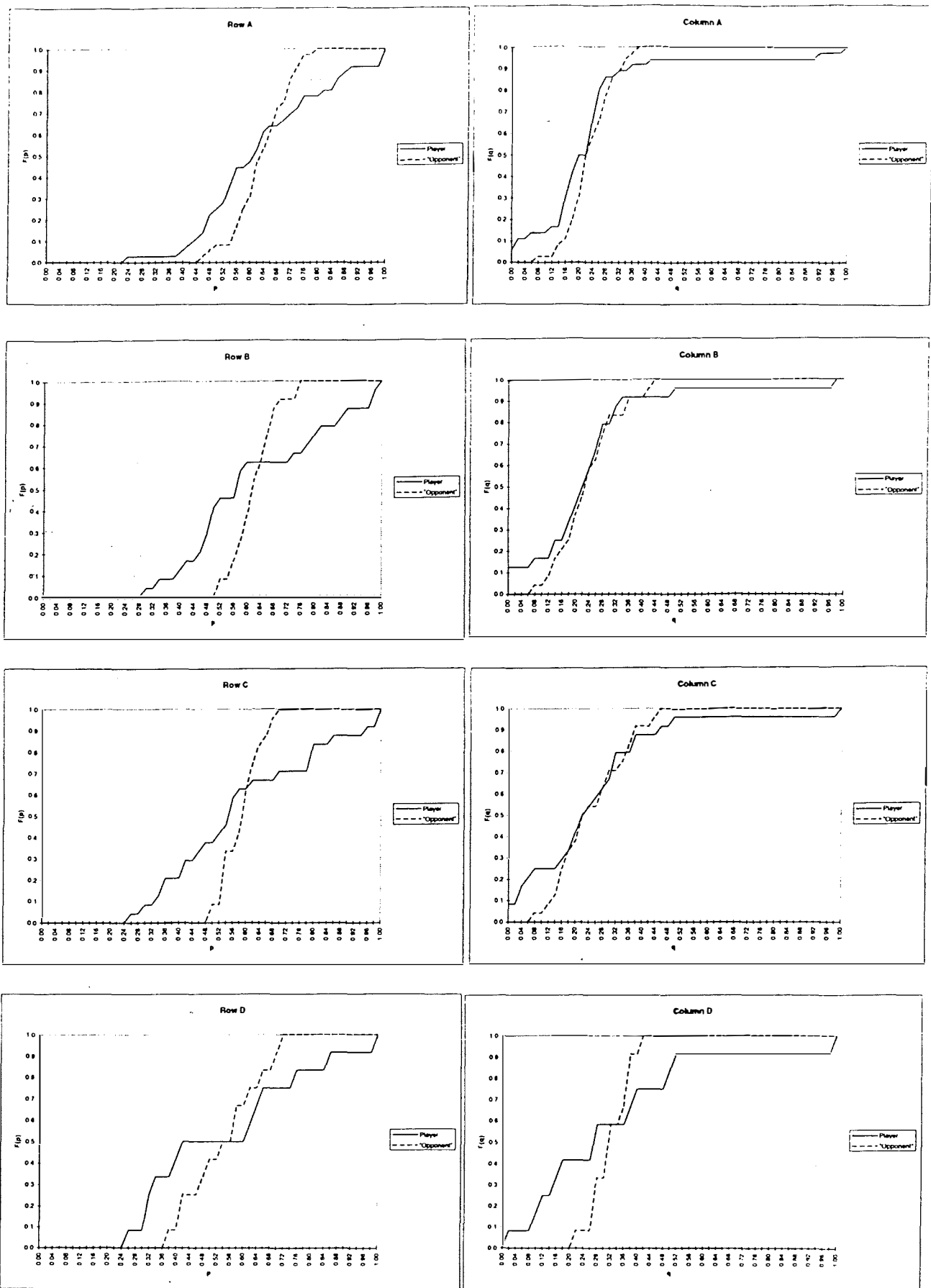


Figure 7: Cumulative density of actual choice frequencies (solid) versus observed frequencies (dotted) in game A

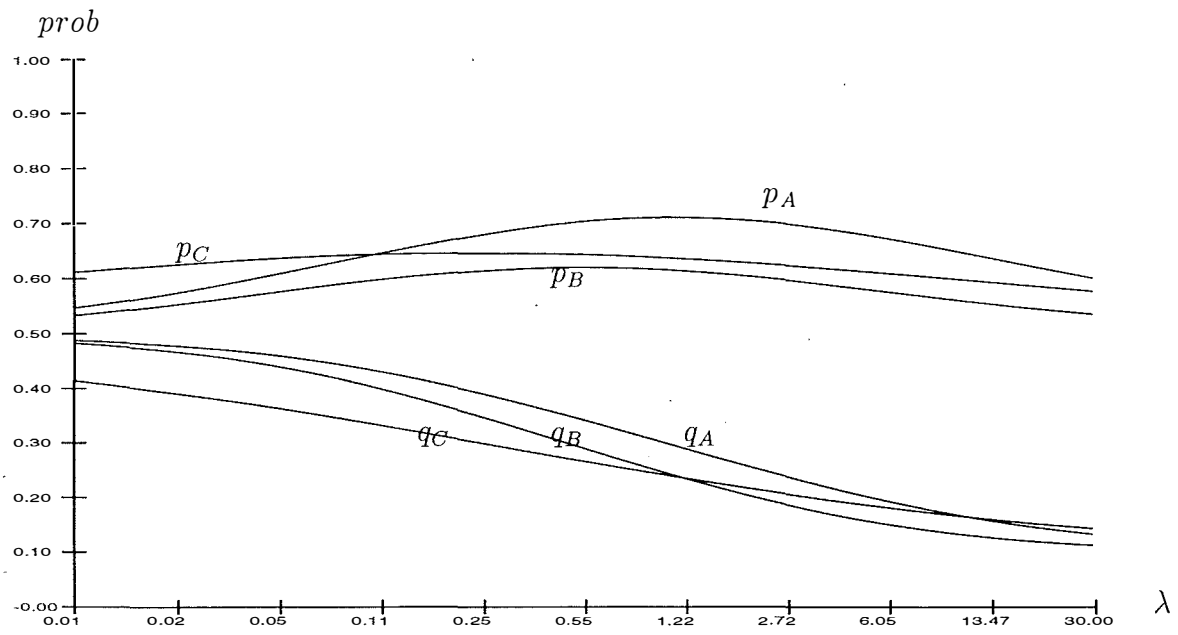


Figure 8: QRE correspondence with heterogeneity for Games A, B, and C

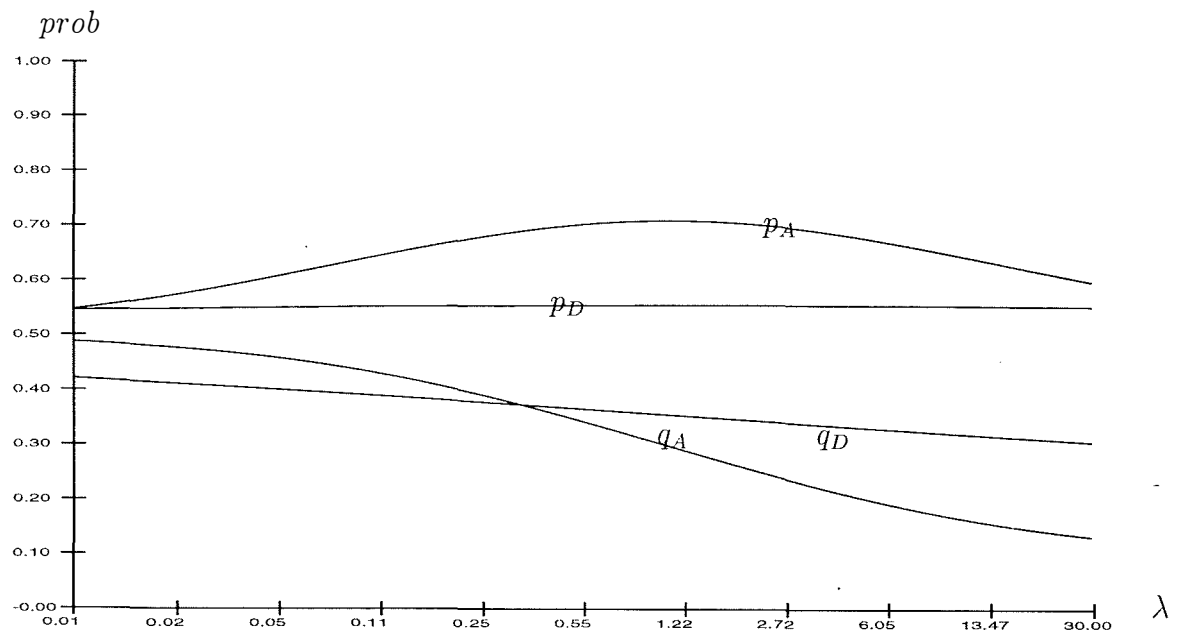


Figure 9: QRE correspondence with heterogeneity for games A and D

striction is that the value of λ is the same for the entire population.

As an initial specification of the distribution of λ_i in the above model, we decided to use a Normal distribution over the values of $\text{Log}(\lambda)$ with mean μ and variance σ^2 .¹⁷ This distribution seems appropriate since it is reasonable to believe that a large percentage of the subjects' behavior is similar and centered around some mean value of $\lambda = e^\mu$.

In order to find the maximum likelihood estimate of the parameters in this model, we conducted a grid search over values of μ and σ . For each individual, the likelihood $L(d_i|\lambda_k)$ of that subject's data d_i was first estimated for each value of λ_k , for $k \in K$, where K is a set of indices, and the λ_k were evenly spaced according to the value of $\text{Log}(\lambda_k)$. Then, for each $k \in K$, a probability weight $\delta_k(\mu, \sigma)$ for λ_k was computed according to the CDF of the normal distribution over $\text{Log}(\lambda_k)$. The likelihood function was computed by first summing across the weighted values of the above likelihoods for all values of $\text{Log}(\lambda)$ and then summing across individuals. The following equation represents the log-likelihood calculation which was performed,

$$\text{Log}L = \sum_{i \in N} \sum_{k \in K} \delta_k(\mu, \sigma) \text{Log}L(d_i|\lambda_k)$$

where N represents the set of subjects for that game.

In order to provide a comparison against which to test the heterogeneous QRE with heterogeneity model, we also incorporated heterogeneity into the NNM in approximately the same way. That is, we incorporated heterogeneity into the error rates of the individuals. Since the NNM error parameter, γ , is only defined over the interval $[0, 1]$ we parameterized the distribution of γ by the Beta distribution with parameters α and β . By searching over the nonnegative range of these parameters, the Beta distribution allows for a flexible estimation of γ -heterogeneity. Similar to the QRE heterogeneity model, the log-likelihood is computed using the following formula:

$$\text{Log}L = \sum_{i \in N} \sum_{j \in J} \delta_j(\alpha, \beta) L(d_i|\gamma_j)$$

where J is the set of indices of γ and δ_j is the probability weight (from the Beta distribution) assigned to a particular value of γ_j .

The results of the grid searches over values of (μ, σ) and (α, β) are reported for each game in Tables 7-10. These estimates produce dramatically different results, compared to the homogeneous error models of QRE and NNM. For QRE, comparing the likelihood values with those from the estimates without heterogeneity, it is apparent that this model

¹⁷Using the Log of λ results in a continuous distribution over $(-\infty, \infty)$ which is necessary for the Normal distribution. This modification is necessary since λ is only specified in the interval $[0, \infty)$.

does a much better job of accurately predicting play, and this improvement is significant for all four games, at any conventional level of significance. The improvement in fit is somewhat less for NNM, but still significant for all games. More importantly, the relative performance of NNM and QRE are reversed. While NNM provided a slightly better fit than QRE for some of the games under the (incorrect) hypothesis of homogeneity, the heterogeneous QRE model fits all the data much better than the heterogeneous NNM model. For all games, the log-likelihood values reported are higher for the QRE model with heterogeneity than for the heterogeneity NNM model. In fact, a more detailed analysis (not reported in the tables here) confirms that this is true for every 10-period subsample.

<i>MODEL</i>	\hat{p}	p	\hat{q}	q	μ	$\lambda = e^\mu$	σ	α	β	$-Log(L)$
QRE	0.682	0.643	0.208	0.241	1.50	4.48	2.2			2012.7
NNM	0.500	0.643	0.215	0.241				1.15	0.50	2164.9

Table 7. Estimates for QRE and NNM Heterogeneity Model for Game A

<i>MODEL</i>	\hat{p}	p	\hat{q}	q	μ	$\lambda = e^\mu$	σ	α	β	$-Log(L)$
QRE	0.611	0.630	0.228	0.244	0.30	1.35	2.0			1367.8
NNM	0.500	0.630	0.230	0.244				1.05	0.50	1449.8

Table 8. Estimates for QRE and NNM Heterogeneity Model for Game B

<i>MODEL</i>	\hat{p}	p	\hat{q}	q	μ	$\lambda = e^\mu$	σ	α	β	$-Log(L)$
QRE	0.637	0.594	0.239	0.257	0.10	1.12	3.6			1351.2
NNM	0.500	0.594	0.230	0.257				0.65	0.30	1451.0

Table 9. Estimates for QRE and NNM Heterogeneity Model for Game C

<i>MODEL</i>	\hat{p}	p	\hat{q}	q	μ	$\lambda = e^\mu$	σ	α	β	$-Log(L)$
QRE	0.555	0.550	0.320	0.328	2.40	11.02	6.0			724.8
NNM	0.500	0.550	0.350	0.328				1.00	0.01	756.7

Table 10. Estimates for QRE and NNM Heterogeneity Model for Game D

The most apparent source of the relative lack of improvement of the NNM fit with heterogeneity is that NNM always predicts the row player will play Up with probability equal to 0.5, for all values of (α, β) . Thus heterogeneity in NNM can only improve the fit to column player data. But introducing heterogeneity into the QRE model reduces the extent to which this model predicts overplaying of U by the row player relative to the experimental results.

Recall that earlier we looked at whether or not the value of λ increased during an experiment, and found no consistent trends. An interesting question to raise, then, is whether the values of μ in the QRE heterogeneity model are increasing across periods. This would again provide a weak test of our hypothesis that the parameter of λ in the QRE is a measure of learning. If the value of μ increases during the experiments, it can be interpreted as evidence of population learning. The results of a test of this hypothesis, however, are again inconclusive. While Kendall rank-order correlation coefficients are positive for Games A, B, and C (respectively, 0.60, 0.95, and 0.20) and significant at $p < 0.05$ for Game B, the coefficient for Game D is negative (-0.67). The same is true for Spearman rank-order coefficients.

6 Conclusion

This paper provided a test of alternative equilibrium notions for a set of 2×2 games with unique mixed strategy equilibria. The results indicate that the heterogeneous error Quantal Response Equilibrium provides the best fit to actual play in these games, compared to Nash equilibrium, random play, and the Noisy Nash Model. These results are supported in tests for all games. However, the data do not always fit the QRE well, since the model generally overpredicts the frequency with which the row player will choose U .

The predictive ability of both QRE and NNM is greatly improved, however, by the introduction of heterogeneity into these models. We model heterogeneity as a distribution over the parameters (λ or γ) in a population. In the QRE model, heterogeneity means that each player has their own value for λ and behaves as if this value represents the parameter for the entire population. The set of these λ_i is represented by a distribution $N(\mu, \sigma^2)$ over the space of $\text{Log}(\lambda)$. A similar construction is done for the NNM using the Beta distribution with parameters α and β .

We also tested whether payoff magnitudes cause significant effects, as predicted by the homogeneous version of QRE if λ is constant across the games being played. This hypothesis is rejected in comparisons between unconstrained estimation of the parameter and constrained estimation, where the parameter is held the same across all the games. The additional hypothesis that the parameter λ represents a measure of learning which increases with play of a game, or exposure to a similar game, also finds little support when comparisons are made between sessions where a game is played first and sessions where it is played second. Both of these results indicate that it would be desirable to try to endogenize λ and to develop a model of how λ changes as players gain experience. This is beyond the scope of the current paper, but McKelvey, Palfrey, and Weber (1997) pursues the first of these two extensions.

Summarizing, our results indicate that, while play differs considerably in these Nash-equivalent games, the differences can be explained somewhat well by a model which introduces both error and heterogeneity within this error.

REFERENCES

- Cooper, R., D. DeJong, R. Forsythe, and T. Ross, "Cooperation Without Reputation: Experimental Evidence from Prisoners' Dilemma Games," *Games and Economic Behavior*, 12:187-218, 1996.
- Cox, J., V. Smith, and J. Walker, "Tests of a Heterogeneous Bidder Theory of First Price Auctions," *Economics Letters*, 12:207-12, 1983.
- El-Gamal, M. and D. Grether, "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, 90:1137-45, 1995.
- El-Gamal, M. and T. Palfrey, "Economical Experiments: Bayesian Efficient Experimental Design," *International Journal of Game Theory*, 1996.
- Forsythe, R., J. Horowitz, N. Savin, and M. Sefton, "Fairness in Simple Bargaining Experiments," *Games and Economic Behavior*, 6:347-69, 1994.
- Grether, D. and Plott, C., "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69:623-38, 1979.
- Harrison, G., "Theory and Misbehavior in First Price Auctions," *American Economic Review*, 79:749-62, 1989.
- Hoffman, E., K. McCabe, and V. Smith, "Expectations and the Monetary Stakes in Ultimatum Games," *International Journal of Game Theory*, 25:289-301, 1996.
- Mookherjee, D. and B. Sopher, "Learning and Decision Costs in Experimental Constant Sum Games," *Games and Economic Behavior*, forthcoming.
- Mookherjee, D. and B. Sopher, "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7:62-91, 1994.
- McKelvey, R. D., and T. Palfrey, "An Experimental Study of the Centipede Game," *Econometrica*, 60:803-36, 1992.
- McKelvey, R. D., and T. R. Palfrey, "Quantal Response Equilibria in Normal Form Games," *Games and Economic Behavior*, 10:6-38, 1995.
- McKelvey, R. D., and T. R. Palfrey, "A Statistical Theory of Equilibrium in Games," *Japanese Economic Review*, 47:186-209, 1996.
- McKelvey, R. D., and T. R. Palfrey, "Quantal Response Equilibria in Extensive Form Games," *Experimental Economics*, 1:9-41, 1998.
- McKelvey, R. D., T. R. Palfrey, and R. A. Weber, "Endogenous Rationality Equilibrium," mimeo, September 1997.

- Luce, D. (1959), *Individual Choice Behavior*, New York: Wiley.
- Ochs, J., "Games with unique mixed strategy equilibria: An experimental study," *Games and Economic Behavior*, 10 (1995):202-217.
- O'Neill, B., "A Nonmetric Test of the Minimax Theory of Two-Person Zerosum Games," *Proceedings of the National Academy of Sciences, USA*, 84:2106-9, 1987.
- Palfrey, T. and J. Prisbrey, "Anomalous Behavior in Public Goods Experiments: How much and Why," *American Economic Review*, 87:829-46, 1997.
- Pommerehne, W., F. Schneider, and P. Zweifel, "Economic Theory of Choice and the Preference Reversal Phenomenon: A Reexamination," *American Economic Review*, 72:569- 74, 1982.
- Rapoport, A. and R. Boebel, "Mixed Strategies in Strictly Competitive Games: A Further Test of the Minimax Hypothesis," *Games and Economic Behavior*, 4:261-83, 1992.
- Siegel, S., "Decision Making and Learning Under Varying Conditions of Reinforcement," *Annals of the New York Academy of Sciences*, 89:766-83, 1961.
- Smith, V. and J. Walker, "Monetary Rewards and Decision Cost in Experimental Economics," *Economic Inquiry*, 31:245-61, 1993.
- Stahl, D. and P. Wilson, "On Players Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10:218-54, 1995.

INSTRUCTIONS

This is an experiment in decision making, and you will be paid for your participation in cash. Different subjects may earn different amounts. What you earn depends partly on your decisions and partly on the decisions of others.

The entire experiment will take place through computer terminals, and all interaction between subjects will take place through the computers. It is important that you not talk or in any way try to communicate with other subjects during the experiment. If you disobey the rules, we will have to ask you to leave the experiment.

We will start with a brief instruction period. During the instruction period, you will be given a complete description of the experiment and will be shown how to use the computers. If you have any questions during the instruction period, raise your hand and your question will be answered so everyone can hear. If any difficulties arise after the experiment has begun, raise your hand, and an experimenter will come and assist you.

The subjects will be divided into two groups, each containing an equal number of subjects. The groups will be labeled the RED group and the BLUE group. To determine which color you are, will you each please select an envelope as the experimenter passes by you.

[EXPERIMENTER PASS OUT ENVELOPES]

Inside each envelope is an index card labeled either BLUE or RED. If you chose BLUE, you will be BLUE for the entire experiment. If you chose RED, you will be RED for the entire experiment. Please remember your color, because the instructions are slightly different for the BLUE and RED subjects.

This experiment will consist of two sessions, Session A and Session B. Each session will consist of several periods or matches. I will now describe what occurs in each match. First, you will be randomly paired with a subject of the opposite color. Thus, if you are a BLUE subject,

you will be paired with a RED subject. If you are a RED subject, you will be paired with a BLUE subject.

[TURN ON OVERHEAD PROJECTOR]

After you have been paired, each subject will simultaneously be asked to make a choice. The RED subject in each pair will be asked to choose one of the two rows in the matrix which will appear on the computer screen and which is also now shown on the screen at the front of the room. The RED subject can choose either “Up” or “Down”. The BLUE subject in each pair will be asked to choose one of the two columns in the matrix, either “Left” or “Right”. Neither subject will be informed of what choice the other subject has made until after all choices have been made.

After each subject has made his or her choice, payoffs for the match are determined based on the choices made. Payoffs to each subject are indicated by the numbers in the matrix. The payoff to the RED subject is in red and appears in the lower left of each compartment, while the payoff to the BLUE subject is in blue and appears in the upper right of each compartment. The units are in dimes.

Thus, if the RED subject chooses “Up” and the BLUE subject chooses “Left” the RED subject receives a payoff of ____ dimes while the BLUE subject receives a payoff of ____ dimes. If the RED subject chooses “Up” and the BLUE subject chooses “Right”, the RED subject receives a payoff of ____ dimes while the BLUE subject receives a payoff of ____ dimes. If the RED subject chooses “Down” and the BLUE subject chooses “Left”, the RED subject receives a payoff of ____ dimes while the BLUE subject receives a payoff of ____ dimes. Lastly, if the RED subject chooses “Down” and the BLUE subject chooses “Right”, the RED subject receives a payoff of ____ dimes while the BLUE subject receives a payoff of ____ dimes.

This process will be repeated for two sessions of ____ matches each. Every match you will be randomly paired with a new subject. You are equally likely to be paired with each subject of the opposite color, however you will never be paired with the same subject twice in a

row. For example, if you are a RED subject, and in match #1 you were paired with BLUE subject #5, then in match #2 you may be paired with any other BLUE subject except BLUE subject #5. The identity of the person you are paired with will never be revealed to you.

Your final earnings for the experiment will be the sum of your payoffs from all ____ matches plus a ____ participation bonus.

[BEGIN COMPUTER INSTRUCTION SESSION]

We will now begin the computer instruction session. Will all RED subjects please move to the computers to my left, near the window, and will all BLUE subjects please move to the computers to my right, near the door to the hallway.

[WAIT FOR SUBJECTS TO MOVE TO APPROPRIATE COMPUTERS]

[TURN OFF OVERHEAD PROJECTOR]

During the computer instruction session, we will teach you how to use the computer by going through a few practice matches. Do not hit any keys until you are told to do so, and when you are told to enter information, type exactly what you are told to type. You are not paid for these practice matches.

Please turn on your computer now by pushing the button labeled "MASTER" on the right hand side of the panel underneath the screen.

[WAIT FOR SUBJECTS TO TURN ON COMPUTER]

When the computer prompts you for your name, type your full name. Then hit the ENTER key. Confirm your entry by pressing the Y key when prompted, or press the N key to correct your entry.

[WAIT FOR SUBJECTS TO ENTER NAMES]

When you are asked to enter your color, type R if your color is RED, and B if your color is BLUE. Then hit ENTER. Confirm your entry by pressing the Y key when prompted, or press the N key to correct your entry.

[WAIT FOR SUBJECTS TO ENTER COLORS]

You now see the experiment screen. Throughout the experiment, you will be told what is currently happening at the left or very bottom of the screen. The strip along the bottom of the screen tells the history of what happened in your previous matches. Since the experiment has not yet begun, this strip along the bottom is currently empty. In the middle of the screen is the matrix which you have previously seen up on the screen at the front of the room. At the top left of the screen you see your color, your subject ID number, and your name. Is there anyone whose color is not correct?

[WAIT FOR RESPONSE]

We will now pass out the experiment record sheet, on which you will record all of the results from this experiment. When you receive an experiment record sheet, please record your name, color, and today's date on the top of the sheet. Do not record your subject ID number at this time.

[EXPERIMENTER PASS OUT EXPERIMENT RECORD SHEETS AND PENCILS]

[WAIT FOR SUBJECTS TO RECORD INFORMATION]

We will now start the first practice match. Remember, do not hit any keys or click the mouse button until you are told to do so.

If you are a RED subject, on the left of the screen you are asked to please choose a row. If you are a BLUE subject, you are asked to please choose a column. You will choose a row or column by moving the mouse to the appropriate choice and clicking the mouse button.

Will all RED subjects now move the mouse so that the arrow on the screen is pointing to the bottom row labeled "D" and will all BLUE subjects now move the mouse so that the arrow on the screen is pointing to the left column labeled "L".

[WAIT FOR SUBJECTS TO MOVE MOUSE TO APPROPRIATE ROW OR COLUMN]

Note that the row or column to which you are pointing with the mouse is now surrounded by a flashing rectangle. Will all RED subjects please choose "Down" and all BLUE subjects please choose "Left" by clicking the mouse button now while the arrow is pointing to the

appropriate row or column. After choosing the row or column, confirm your choice by clicking on the “Yes” icon at the bottom of the screen or click on the “No” icon to correct your choice.

[WAIT FOR SUBJECTS TO CHOOSE ROW OR COLUMN AND CONFIRM THE CHOICE]

After all subjects have confirmed their choices, the match is over. The outcome of this match, Down-Left, is now highlighted on everybody’s screen. Also note that the moves and payoffs of the match are recorded in the experiment history at the bottom of the screen. The outcomes of all of your previous matches will be recorded at the bottom of the screen throughout the experiment so that you can refer back to previous outcomes whenever you like. The payoff to the RED subject for this match is ____ and the payoff to the BLUE subject is _____. Please record the outcome of this match on your experiment record sheet in the first row labeled “PRACTICE”. After you have finished recording the outcome of this match, use the mouse to click on the “OK” icon at the bottom of the screen to indicate that you are ready to continue.

[WAIT FOR SUBJECTS TO RECORD OUTCOME AND CLICK “OK”]

You are not being paid for the practice session, but if this were the real experiment, then the payoff you have recorded would be money you have earned from the first match, and you would be paid this amount for that match at the end of the experiment. The total you earn over all ____ real matches, in addition to the _____ bonus, is what you will be paid for your participation in the experiment.

We will now proceed to the second practice match.

[EXPERIMENTER HIT KEY TO START SECOND MATCH]

For the second match, each subject has been randomly paired with a different subject of the opposite color. You are not paired with the same subject you were paired with in the first match. The rules for the second match are exactly like the first. Will all RED subjects again choose “Down” by clicking on the bottom row and confirming the choice. Also, will all BLUE subjects choose “Right” by clicking on the right column and confirming the choice.

[WAIT FOR SUBJECTS TO CHOOSE ROW OR COLUMN AND CONFIRM CHOICE]

The outcome of this match, Down-Right, is now highlighted on everybody's screen. The payoff to the RED subject for this match is ____ and the payoff to the BLUE subject is _____. Please record the outcome of this match on your experiment record sheet in the second row labeled "PRACTICE". After you have finished recording the outcome of this match, use the mouse to click on the "OK" icon at the bottom of the screen to indicate that you are ready to continue.

[WAIT FOR SUBJECTS TO RECORD OUTCOME AND CLICK "OK"]

We will now proceed to the third practice match.

[EXPERIMENTER HIT KEY TO START THIRD MATCH]

Will all RED subjects choose "Up" by clicking on the top row and confirming the choice, and will all BLUE subjects choose "Left" by clicking on the left column and confirming the choice.

[WAIT FOR SUBJECTS TO CHOOSE ROW OR COLUMN AND CONFIRM CHOICE]

The outcome of this match is now highlighted on everybody's screen. The payoff to the RED subject for this match is ____ and the payoff to the BLUE subject is _____. Please record the outcome of this match on your experiment record sheet in the third row labeled "PRACTICE". After you have finished recording the outcome of this match, use the mouse to click on the "OK" icon at the bottom of the screen to indicate that you are ready to continue.

[WAIT FOR SUBJECTS TO RECORD OUTCOME AND CLICK "OK"]

We will now proceed to the fourth practice match.

[EXPERIMENTER HIT KEY TO START FOURTH MATCH]

Will all RED subjects choose "Up" by clicking on the top row and confirming the choice, and will all BLUE subjects choose "Right" by clicking on the right column and confirming the choice.

[WAIT FOR SUBJECTS TO CHOOSE ROW OR COLUMN AND CONFIRM THE CHOICE]

The outcome of this match is now highlighted on everybody's screen. The payoff to the RED subject for this match is ____ and the payoff to the BLUE subject is _____. Please record

the outcome of this match on your experiment record sheet in the fourth row labeled "PRACTICE". After you have finished recording the outcome of this match, use the mouse to click on the "OK" icon at the bottom of the screen to indicate that you are ready to continue.

[WAIT FOR SUBJECTS TO RECORD OUTCOME AND CLICK "OK"]

[EXPERIMENTER HIT KEY TO END PRACTICE SESSION]

This concludes the practice matches. The computer screen now indicates your total payoff for the four practice matches. This is the amount you would have earned for these matches if these were matches in the actual experiment. You do not need to record this total.

In the actual experiment there will be two sessions, Session A and Session B, of ____ matches each, and, of course, it will be up to you to make your own decisions. At the end of the Session B, the experiment ends and we will pay each of you privately, in cash, the total amount you have accumulated during all ____ matches, plus your guaranteed ____ participation bonus. No other person will be told how much cash you earned in the experiment. You need not tell any other participants how much you earned.

Are there any questions before we begin Session A?

[EXPERIMENTER TAKE QUESTIONS]

O.K., then we will now begin with the actual experiment and Session A. Please press the spacebar once and wait a moment for the current screen to clear.

[WAIT FOR SUBJECTS TO PRESS SPACEBAR AND CLEAR SCREEN]

After the screen has changed, please type "DL" and hit the "Enter" Key.

[EXPERIMENTER START EXPERIMENT PROGRAM]

If there are any problems from this point on, raise your hand and an experimenter will come and assist you. When the computer asks for your name, please start as before by typing your name. Wait for the computer to ask for your color, then respond with the correct color.

[WAIT FOR SUBJECTS TO INPUT NAME AND COLOR]

[EXPERIMENTER HIT KEY TO START MATCHING]

At the top left of your screen you once again see your color and subject ID number. This ID number may be different than the ID number you were given during the practice matches. Will you now please record this ID number at the top of your experiment record sheet where it says "Subject ID#, Session A". Also, please make sure that the color indicated on the screen is correct.

[WAIT FOR SUBJECTS TO RECORD ID NUMBERS]

Okay, we will now begin Session A.

[START EXPERIMENT]

[AFTER 1ST MATCH, REMIND SUBJECTS THEY ARE PAIRED WITH A NEW PERSON]

2X2 EXPERIMENT

INTER-SESSION INSTRUCTIONS

[AFTER MATCH # ____, EXPERIMENTER TERMINATE EXPERIMENT]

Session A is now completed. Please record your total payoff for Session A by adding up all your payoffs for matches 1 through ____ and writing this total at the bottom of your experiment record sheet where it says "Session A Total". Make sure that you do not include your payoffs from the practice sessions when adding up your payoffs. Your total payoff should match the total that is indicated on the computer screen. If there are any problems, please raise your hand.

[WAIT FOR SUBJECTS TO COMPUTE THEIR SESSION A TOTAL]

[TURN ON OVERHEAD PROJECTOR WITH NEW MATRIX]

In a minute we will begin Session B. During Session B, all the rules will be the same, however the payoffs for each match will change slightly. In Session B, if the RED subject chooses "Up" and the BLUE subject chooses "Left" the RED subject receives a payoff of ____ dimes and the BLUE subject receives a payoff of ____ dimes. If the RED subject chooses "Up" and the BLUE subject chooses "Right", the RED subject now receives a payoff of ____ dimes and the BLUE subject receives a payoff of ____ dimes. If the RED subject chooses "Down" and the BLUE subject chooses "Left" the RED subject receives a payoff of ____ dimes and the BLUE subject receives a payoff of ____ dimes. Lastly, if the RED subject chooses "Down" and the BLUE subject chooses "Right" the RED subject receives a payoff of ____ dimes and the BLUE subject receives a payoff of ____ dimes.

Are there any questions before we begin Session B?

[EXPERIMENTER TAKE QUESTIONS]

O.K., then we will now begin Session B. Please press the spacebar once and wait a moment for the current screen to clear.

[WAIT FOR SUBJECTS TO PRESS SPACEBAR AND CLEAR SCREEN]

After the screen has changed, please type "DL" and hit the "Enter" Key.

[EXPERIMENTER START EXPERIMENT PROGRAM]

[TURN OFF OVERHEAD PROJECTOR]

When the computer asks for your name, please start as before by typing your name. Wait for the computer to ask for your color, then respond with the correct color.

[WAIT FOR SUBJECTS TO INPUT NAME AND COLOR]

[EXPERIMENTER HIT KEY TO START MATCHING]

At the top left of your screen you once again see your color and subject ID number. This ID number may be different than the ID number you were given previously. Will you now please record this ID number at the top of your experiment record sheet where it says "Subject ID#, Session B". Also, please make sure that the color indicated on the screen is correct.

[WAIT FOR SUBJECTS TO RECORD ID NUMBERS]

Okay, we will now begin Session B.

[START EXPERIMENT]

[AFTER 1ST MATCH, REMIND SUBJECTS THEY ARE PAIRED WITH A NEW PERSON]

FINAL INSTRUCTIONS

The experiment is now completed. Please record your total payoff for Session B by adding up all your payoffs for matches 1 through ____ and writing this total at the bottom of your experiment record sheet where it says "Session B Total". Your total payoff for Session B should match the total that is indicated on the computer screen.

After you have calculated this payoff, add together your Session A total, Session B total, and the ____ participation bonus to get your final payoff for the experiment. Remember that the payoff units are in ____.

Also, at the bottom of the last page of your record sheet you are asked to write your name, social security number, amount received, and your signature. You may write your name and social security number now, however please wait until after you have received payment to write the amount received and your signature. If there are any problems or questions, please raise your hand.

After you are done calculating your payoff for the experiment, please remain seated. You will be paid in the office at the back of the room one at a time.

Please take a look at the index card labeled "BLUE" or "RED" which you received at the beginning of the experiment. The number in the bottom right-hand corner of this card will determine the order in which you will be paid. When we call your number, it is your turn to receive payment.

Please bring all your things with you when you go to the back office. You can leave the experiment through the back door of the office.

Please refrain from discussing this experiment while you are waiting to receive payment so that privacy regarding individual choices and payoffs may be maintained. While you are waiting I will come by to gather experiment materials. When you are done calculating your payoff for the experiment, you may turn off your computer by pushing the button labeled "MASTER" on the right hand side of the panel underneath the screen.

Whenever the person with the index card labeled number one is ready, he or she may now go to the back office to receive payment.

Thank you all very much for participating in this experiment.